Gender and Age Estimation Using Facial Gestures

Ege SARIÇA Hasan Emre KAYIR

***Abstract*—This project focuses on gender and age estimation using facial gestures. The dataset consists of two subsets: the first subset contains around 600 images from various age groups to test the algorithm's basic functionality, while the second subset comprises 2448 images of individuals aged 15-65, including variations in image quality and demographics. The mean absolute error for age estimation increases from 3.11 to 5.77 when low-quality images are introduced, but remains sufficient for the project's objectives. Facial gestures are utilized for feature extraction, with Histogram of Oriented Gradients (HOG) proving to be the most effective feature. The HOG technique captures gradient orientation occurrences in localized portions of the image. To handle the large number of features, Principle Component Analysis (PCA) is applied for feature reduction. For gender estimation, a Support Vector Machine (SVM) classifier is trained using the HOG feature vectors, while age estimation utilizes a Random Forest regressor. The trained models achieve satisfactory results in predicting gender and age from facial images.**

I. Dataset

Dataset to use during training and testing process has huge importance for quality and performance results of the project. 2 different subsets of UTKFaces [1] dataset are used. First subset contains around 600 images from various age groups. These images have good quality in terms of contrast, resolution, clearness and face gestures. Reason for constructing this subset is to test the basic working process and outputs of the algorithm. Second subset contains 24 male and 24 female from each age between 15-65 which is our target group, totally 2448 images. These images have different contrast, resolution, blurriness, race, creating overall various quality images. The reason for constructing this subset is to test the performance of our algorithm when bad quality image is given. Results conclude that mean absolute error increase from 3.11 to 5.77 for age estimation, which is sufficient for our project's purpose. Mean absolute error of 5.77 is mostly consist of the prediction error for age groups at edges.

II. Facial Features

Many facial features can be used for gender and age estimation: Gradient, Hessian, Histogram of the Oriented Gradients, ratios of face parts, wrinkles, hair, beard, moustache. [2] After examining the sample projects, we have seen that best result is achieved by using Histogram of Oriented Gradients. Other features may even cause more error when added as second feature. That's why HOG usage is focused on. HOG counts occurrences of gradient orientation in localized portions of an image. For 200x200 image of our case:

- Cell size 8x8 = 25x25 = 625 cells
- Block size 2x2 = 24x24 = 576 blocks
- Bins (9) = histogram of gradients are created with 9 bins
- Each block has, 4 cells x 9 bins = 36 features, 576 x 36 = 20736 features create an output vector

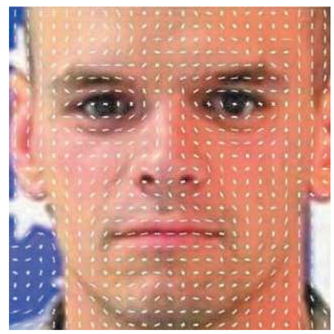

Fig. 1. An example of HOG of an facial image

Since 20736 features is too much for algorithm to process, feature reduction technique, PCA is used.

III. Principle Component Analysis (PCA)

For large data of 2448 image, 2448x20736 size matrix is given as input to ML algorithm. This results in the program to work unacceptably slow. Instead, a feature reduction technique, PCA is used. [3] Basically, PCA algorithm takes 2448x20736 size HOG output, creates principal components, computes variance of each principal component and eliminates the components with low variance, i.e. Gradient magnitude and phase components which are similar to each other. We have taken into account first 100 highest variance components. After applying PCA, mean absolute error almost stays the same. Working time of the project is yet dropped from 101 seconds to 22.7 seconds, which is a massive improvement for the algorithm.

IV. Gender Estimation

The feature vectors obtained using the HOG technique were utilized to train a Support Vector Machine (SVM) classifier. SVM is a widely used supervised learning algorithm that aims to find an optimal hyperplane to separate data points of different classes. It works by mapping the feature vectors into a higher-dimensional space and maximizing the margin between classes. [4]

To train the SVM classifier for gender estimation, we utilized the labeled images from our dataset. The feature vectors, derived through the HOG feature extraction process, were used as input to the SVM algorithm. The SVM classifier learned the underlying patterns and relationships between the HOG features and gender labels during the training phase.

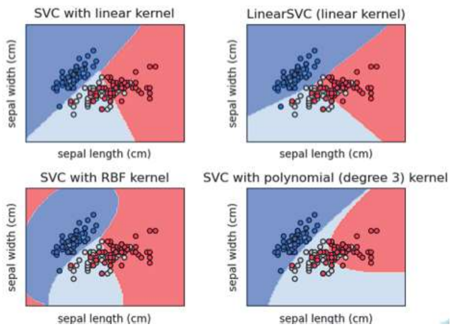

Fig. 2. A simple representation of SVM with different types of kernel functions [5]

Once the SVM classifier was trained, we applied it to estimate the gender of test images. For a given test image, we followed a similar process of feature extraction using the HOG technique. The resulting feature vector was then fed into the trained SVM classifier.

The SVM classifier utilized the learned patterns to predict the gender of the test image. The output of the classifier was either ”0” (male) or ”1” (female), representing the estimated gender based on the input features.

V. Age Regression

The feature vectors obtained through the HOG technique were also used to train a Random Forest regressor. Random Forest is an ensemble learning algorithm that combines multiple decision trees to make predictions. It is well-suited for regression tasks as it can capture non-linear relationships between the input features and the target variable. [6]

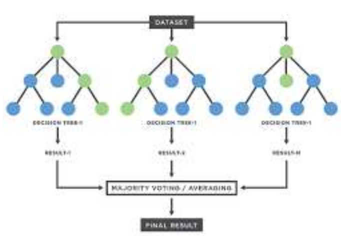

Fig. 3. A representation of Random Forest [7]

To train the Random Forest regressor for age estimation, we utilized the labeled images from our dataset. The feature vectors, extracted using the HOG technique, were used as input to the Random Forest algorithm. Each decision tree in the “forest” is trained by different random subsets of these feature vectors. The regressor learned the underlying patterns and relationships between the HOG features and the age labels during the training phase. When making a prediction, every tree in the forest makes its own prediction and the final prediction is based on the average of all the individual predictions.

VI. Results

A. Age Estimation

For age estimation, we evaluated the performance of our system using mean absolute error (MAE) as the evaluation metric. We compared the results obtained without Principal Component Analysis (PCA) and with PCA. The results are as follows:

a) Mean Absolute Error without PCA: 5.77:

- This indicates that, on average, the predicted age differed from the actual age by approximately 5.77 years.
- The age estimation model achieved this result without employing PCA.
- The working time required for age estimation was approximately 101 seconds.

b) Mean Absolute Error with PCA: 5.72:

- By incorporating PCA into our age estimation pipeline, we achieved a slight improvement in accuracy.
- The mean absolute error reduced to 5.72 years, indicating a slightly better estimation performance.
- The working time for age estimation with PCA was significantly reduced to approximately 11.9 seconds.

These results demonstrate the effectiveness of our age estimation model in predicting the age of individuals from face images. The achieved mean absolute errors, both with and without PCA, indicate a reasonable level of accuracy in estimating age. The inclusion of PCA led to a slight improvement in accuracy while significantly reducing the processing time, making it a valuable addition to our approach. Overall, these results are satisfactory, aligning with our project goal.

B. Gender Estimation

For gender estimation, we evaluated the performance of our system using accuracy as the evaluation metric. The results are as follows:

a) Accuracy: 95.05%:

- The gender estimation model achieved an accuracy of 95.05% in correctly identifying the gender of individuals from their face images.
- This high accuracy highlights the effectiveness of the SVM algorithm in capturing the distinguishing HOG features between male and female faces.
- The working time required for gender estimation was approximately 9.7 seconds.

The results of our gender estimation model demonstrate its strong performance in accurately identifying the gender of individuals from face images. With a high accuracy of 95.05%, our model successfully achieves our project goal of gender estimation. These results are highly satisfactory and confirm the effectiveness of our approach.

REFERENCES

[1] https://susanqq.github.io/UTKFace/

[2] B. Bas, J. Ahlberg, and R. Forchheimer, Facial Image-Based Gender and Age Estimation, Linköping University, p. 9, August 2013.

[3] M. G. Sarowar, M. A. Razzak and M. A. A. Fuad, ”HOG feature descriptor based PCA with SVM for efficient & accurate classification of objects in image,” 2019 IEEE 9th International Conference on Advanced Computing (IACC), Tiruchirappalli, India, 2019, pp. 171-175

[4] Yinglin Xia, Progress in Molecular Biology and Translational Science, Academic Press, Volume 171, 2020, pp. 309-491

[5] https://scikit-learn.org/stable/modules/svm.html

[6] Breiman, L. (2001). Random Forests. Machine Learning. 45. 5-32.

[7] https://www.tibco.com/reference-center/what-is-a-random-forest